

Gráficos estadísticos

Francisco García
(fgarcia@cipf.es)

CIPF's Research Development Programme

29 de octubre de 2014

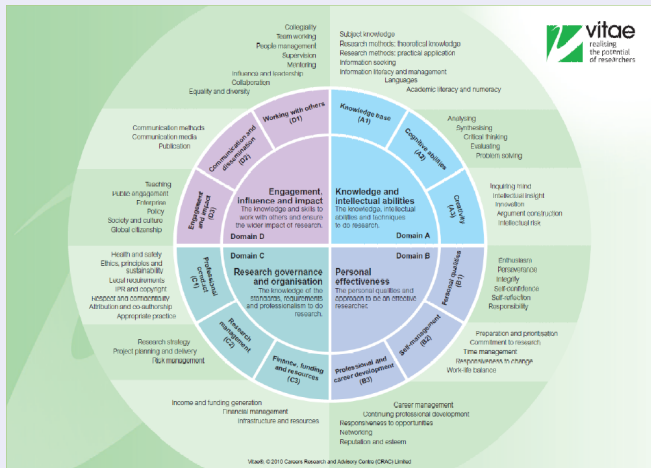
Índice

- 1 Introducción
- 2 Gráficos con R
 - ¿Qué es **R**?
 - Ejemplos y ejercicios
- 3 Herramientas web
 - Paintomics
 - Babelomics
 - CellMaps

Índice

- 1 Introducción
- 2 Gráficos con R
 - ¿Qué es **R**?
 - Ejemplos y ejercicios
- 3 Herramientas web
 - Paintomics
 - Babelomics
 - CellMaps

CIPF's Research Development Programme



Gráficos en publicaciones, informes, presentaciones...



¿Qué recursos utilizaremos para generar gráficos?

- Excel, Power-Point.... no las utilizaremos porque ya las conocéis ;)
- **R** es una potente **herramienta de análisis estadístico y producción de gráficos**.
- **Herramientas web** que representan gráficamente los resultados obtenidos en distintos análisis estadísticos.
- Esta sesión está relacionada con la que ofrecerá Ibo Galindo sobre la **imagen digital en investigación y publicación científica**.

Índice

- 1 Introducción
- 2 Gráficos con R
 - ¿Qué es **R**?
 - Ejemplos y ejercicios
- 3 Herramientas web
 - Paintomics
 - Babelomics
 - CellMaps

¿Qué es R?

- R es un entorno de programación que permite hacer análisis estadísticos.
- Es una potente herramienta para generar gráficos de cualquier tipo.
- R es de código abierto y pertenece al proyecto GNU de software libre.
- Disponible para las plataformas Linux, Macintosh y Windows.
- R es un software orientado tanto a usuarios principiantes como profesionales e investigadores que necesiten analizar y representar gráficamente sus datos.
- R es gratuito. No necesitamos ningún tipo de licencia.

¿Cómo obtenemos e instalamos **R**?

- Descargamos **R** en <http://cran.r-project.org/>
- Tras la descarga, ejecutamos el archivo y aparecerá un asistente que nos guiará en el proceso. En unos minutos el software quedará instalado.
- Va bien instalarse **RStudio**, un interfaz que facilita el trabajo con **R**. Está disponible en <http://www.rstudio.com/>
- También hay disponible una versión libre. El modo de instalación es similar: descargamos la herramienta y la instalamos siguiendo el asistente.
- De modo que podemos trabajar directamente desde **R** o bien desde **RStudio**.

Vamos a por la primera sesión de R

- Abrimos **R**
- También abrimos este documento con ejemplos y ejercicios: **enlace**
- Para cada gráfico ejecutamos sus respectivos comandos e intentamos asociar qué es lo que hace cada uno de ellos.
- Al final del documento, hay unos ejercicios que nos están esperando.

Índice

- 1 Introducción
- 2 Gráficos con R
 - ¿Qué es **R**?
 - Ejemplos y ejercicios
- 3 **Herramientas web**
 - Paintomics
 - Babelomics
 - CellMaps

Herramientas web

- Nos ayudan a representar y visualizar los resultados de un análisis estadístico en un entorno biológico.
- Sólo necesitamos un navegador web para acceder a la herramienta y nuestros datos.
- Veremos algunas ideas básicas sobre Paintomics, Babelomics, CellMaps.

Paintomics

- Es una herramienta de integración y visualización de datos ómicos.
- El **input** es el resultado que hemos obtenido en un análisis de datos transcriptómicos y metabolómicos.
- El **output** es la representación gráfica de pathways de señalización (**KEGG**), incluyendo los datos proporcionados anteriormente.
- **Tutorial:** <http://www.paintomics.org/userguide/index.html>

Input

name	fc_g1g2	fc_g1g3
DCP1A	0.359	-0.286
A_32_P205522	0.867	-0.055
CNOT4	0.550	-0.009
DIS3	0.271	-0.291
EXOSC3	0.182	-0.083
EXOSC1	0.021	-0.018
ZCCHC7	0.193	0.040
LSM7	0.120	0.650
PAPD7	0.150	-0.227
A_23_P113263	0.280	-0.605
XRN1	0.209	0.301
DCP2	0.157	0.515
HSPA9	0.271	0.208
ENO3	0.103	-0.180
CNOT8	0.075	-0.005
CNOT7	0.206	0.123

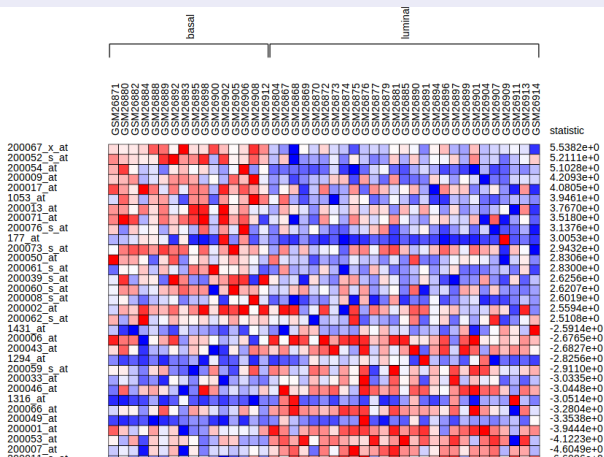
Babelomics

- Es una herramienta de análisis de datos transcriptómicos, proteómicos,... También para cualquier grupo de datos biológicos o clínicos.
- **Input:** son ficheros de texto que incluyen datos en formato rectangular o matricial.
- **Output:** gráficos: heatmaps, árboles de clustering, gráficos de análisis de componentes principales,...
- **Tutorial:** <http://bioinfo.cipf.es/babelomicstutorial/>

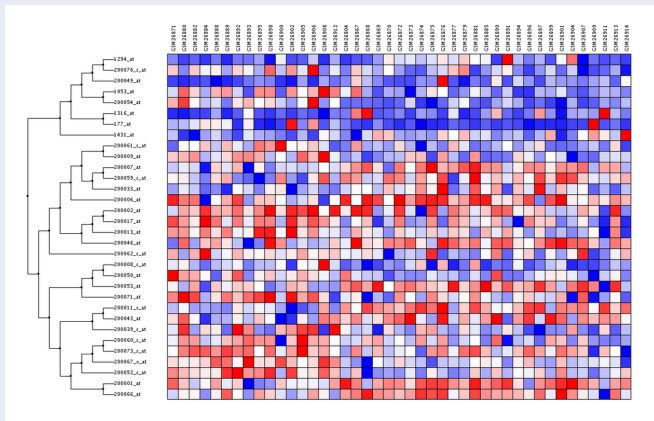
Input

#VARIABLE	BASAL	BASAL	BASAL	LUMINAL	LUMINAL
#NAMES	muestra1	muestra2	muestra3	muestra4	muestra5
gen1	6.88	7.53	7.23	7.16	7.37
gen2	7.39	8.38	8.01	7.94	8.19
gen3	6.65	7.31	7.15	7.06	7.24
gen4	8.91	9.76	9.39	9.30	9.60
gen5	8.59	9.45	9.09	9.08	9.34

Output: heatmap con diferencias entre grupos



Output: clustering de individuos



Vamos a trabajar con la herramienta: **Babelomics**

CellMaps

- Es una herramienta que permite la integración, visualización y el análisis de **redes biológicas**.
- El **input** es un fichero donde indicamos las relaciones entre los nodos de nuestra red. Opcionalmente podemos incluir un fichero con los **atributos** de cada nodo.
- El **output** gráfico es una red en la que se muestran las relaciones de los distintos nodos que la integran. Simultáneamente es posible visualizar los atributos de los elementos biológicos de la red.
- **Tutorial:** <https://github.com/opencb/cell-maps/wiki>

Input

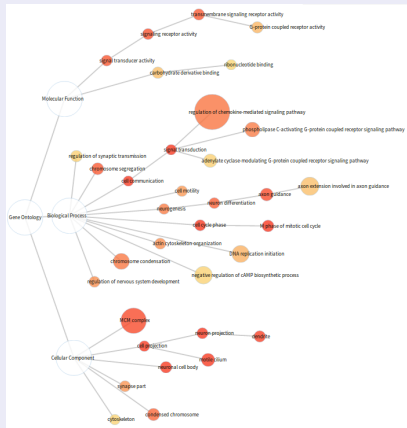
```

GO:0000001> pp> GO:0003674
GO:0000001> pp> GO:0005575
GO:0000001> pp> GO:0008150
GO:0003674> pp> GO:0004871
GO:0004871> pp> GO:0038023
GO:0038023> pp> GO:0004888
GO:0004888> pp> GO:0004930
GO:0003674> pp> GO:0097367
GO:0097367> pp>
GO:0005575> pp>
GO:0005575> pp>
GO:0005575> pp>
GO:0005575> pp>
GO:0042995> pp>
GO:0043005> pp>
GO:0042995> pp>
GO:0005575> pp>

```

ID	pvalor	ind2	descriptor
GO:0031514	0.001	0.16	motile cilium
GO:0000793	0.013	0.129	condensed chromosome
GO:0043025	0.001	0.1	neuronal cell body
GO:0030425	0.003	0.094	dendrite
GO:0044456	0.026	0.086	synapse part
GO:0043005	0.000	0.08	neuron projection
GO:0042995	0.001	0.067	cell projection
GO:0005856	0.044	0.059	cytoskeleton

Output



Vamos a trabajar con la herramienta: **Cellmaps**

